Development and validation of a microRNA-based signature (MiROvaR) to predict early relapse or progression of epithelial ovarian cancer: a cohort study

MiROvaR, a microRNA-based panel to predict early relapse/progression of ovarian cancer patients: a Multicenter Italian Trial in Ovarian (MITO) retrospective translational study.

Marina Bagnoli PhD[a§], Silvana Canevari PhD[b§], Daniela Califano PhD[c], Simona Losito MD[d], Prof. Massimo Di Maio MD[e], Francesco Raspagliesi MD[f], Maria Luisa Carcangiu MD[g], Giuseppe Toffoli MD[h], Erika Cecchin PhD[h], Roberto Sorio MD[i], Vincenzo Canzonieri MD[j], Daniela Russo PhD[c], Giosué Scognamiglio PhD[d], Gennaro Chiappetta PhD[c], Gustavo Baldassarre MD[k], Domenica Lorusso MD[f], Prof. Giovanni Scambia MD[l], Prof. Gian Franco Zannoni MD[m], Antonella Savarese MD[n], Mariantonia Carosi MD[o], Paolo Scollo MD[p], Enrico Breda MD[q], Viviana Murgia MD[r], Francesco Perrone MD[e], Sandro Pignata MD[s#], Loris De Cecco PhD[b#], Delia Mezzanzanica PhD [a#$]

a.Molecular Therapies Unit, Dept of Experimental Oncology and Molecular Medicine, Fondazione IRCCS Istituto Nazionale dei Tumori Milan, Italy (M. Bagnoli PhD, D. Mezzanzanica PhD);

b. Functional Genomics, Dept of Experimental Oncology and Molecular Medicine, Fondazione IRCCS Istituto Nazionale dei Tumori Milan, Italy (S. Canevari PhD, L. De Cecco PhD)**;**

c. Functional Genomic Unit, Istituto Nazionale per lo Studio e la Cura dei Tumori "Fondazione G. Pascale", IRCCS, Naples, Italy, (D. Califano, PhD, G. Chiappetta PhD, D. Russo PhD);

d. Surgical Pathology Unit, Istituto Nazionale per lo Studio e la Cura dei Tumori "Fondazione G. Pascale", IRCCS, Naples, Italy, (S. Losito MD, G. Scognamiglio PhD);

e. Clinical Trials Unit, Istituto Nazionale per lo Studio e la Cura dei Tumori "Fondazione G. Pascale", IRCCS, Naples, Italy, [M. Di Maio MD (present address: Oncology Dept, Università degli Studi di Torino, Ospedale Mauriziano, Torino Italy), F. Perrone, MD];

f. Unit of Gynaecological Oncology, Fondazione IRCCS Istituto Nazionale dei Tumori Milan, Italy (F. Raspagliesi MD, D. Lorusso MD);

g. Anatomic Pathology 1 Unit, Fondazione IRCCS Istituto Nazionale dei Tumori Milan, Italy (ML Carcangiu MD);

h. Experimental and Clinical Pharmacology Unit, Centro di Riferimento Oncologico, Istituto Ricovero e Cura Carattere Scientifico (CRO-IRCCS), Aviano, Italy (G. Toffoli MD, E. Cecchin, PhD);

i. Medical Oncology C, Centro di Riferimento Oncologico, Istituto Ricovero e Cura Carattere Scientifico (CRO-IRCCS), Aviano, Italy, (R. Sorio, MD);

j. Unit of Pathology, Centro di Riferimento Oncologico, Istituto Ricovero e Cura Carattere Scientifico (CRO-IRCCS), Aviano, Italy, (V. Canzonieri, MD);

k. Division of Experimental Oncology 2, Centro di Riferimento Oncologico, Istituto Ricovero e Cura Carattere Scientifico (CRO-IRCCS), Aviano, Italy, (G. Baldassarre, MD);

l. Department of Obstetrics and Gynecology, Gynecologic Oncology Unit, Catholic University of the Sacred Heart, Rome, Italy (G. Scambia MD);

m. Department of Human Pathology, Division of Gynecologic Pathology, Catholic University of the Sacred Heart, Rome, Italy (GF. Zannoni MD);

n. Division of Medical Oncology 1, Regina Elena Cancer Institute, Rome, Italy (A. Savarese, MD);

o. Division of Pathology Regina Elena Cancer Institute, Rome, Italy (M. Carosi, MD);

p. Department of Obstetrics and Gynecology, Azienda Ospedaliera Cannizzaro, Catania (P. Scollo MD);

q. Medical Oncology Unit Ospedale S Giovanni Calibita Fatebenefratelli, Rome, Italy (E. Breda MD)

r. Medical Oncology Unit Ospedale S. Chiara, Trento, Italy (V. Murgia MD);

s. Dept of Urogynaecological Oncology, Istituto Nazionale per lo Studio e la Cura dei Tumori "Fondazione G. Pascale", IRCCS, Naples, Italy, (S. Pignata MD).

[§] Equally contributing first authors

[#] Equally contributing last authors

[$] Corresponding author:
Delia Mezzanzanica - Unit of Molecular Therapies, Department of Experimental Oncology and Molecular Medicine - Fondazione IRCCS Istituto Nazionale dei Tumori, Via Amadeo, 42; 20133 Milan Italy. delia.mezzanzanica@istitutotumori.mi.it. Phone: +390223902569

Keywords: Ovarian cancer; microRNA; molecular predictor; MITO translational group; early relapse; miR-506 family; miR-200 family.

SUMMARY

Background

Despite the efficacy of first-line treatment, risk of relapse/progression remains a challenge for most epithelial ovarian cancer (EOC) patients and development of a molecular predictor could be a valuable tool for patients' stratification. Considering that a master layer of gene expression regulation is provided by microRNAs (miRNAs), we aimed to develop a miRNA-based molecular classifier able to predict progression in EOC patients.

Methods

We analysed miRNA expression profiles in three case materials collected at diagnosis: 179 samples from a MITO (Multicenter Italian Trial in Ovarian cancer) trial (OC179) were used as training set; 263 samples from two of our centres (OC263) and 452 samples from The Cancer Genome Atlas EOC series (OC452), were used as validation sets. We defined progression-free survival (PFS) as the primary clinical endpoint and adapted a semi-supervised prediction method to the miRNA expression profile to define the risk of progression. The predictor's prognostic impact was evaluated by multivariable analysis using a Cox regression model.

Findings

We developed a 35 miRNAs-based predictor of Risk of Ovarian Cancer Relapse/progression (MiROvaR) able to classify OC179 patients into high-risk (89 patients, median PFS 18 months) and low-risk (90 patients, median PFS 38 months) (HR 1·85, 95% CI 1·29–2·60, P<0·001). MiROvaR prognostic value was also significant in the two validation sets and it maintained independent prognostic impact in multivariable analyses adjusting for relevant clinical covariates (HR 1·48, 95% CI 1·03–2·1; P=0·036; HR 3·09, 95% CI 2·23–4·28, P<0·0001; and HR 1·41, 95% CI 1·11–1·79, P=0·0047 for OC179, OC263, and OC452, respectively). MiROvaR performance was confirmed in all Type-II and in the subset of high-grade serous cases present in the OC263 samples thus supporting its value in stratifying patients according to risk of progression regardless of the clinical–pathological characteristics of the tumours at presentation.

Interpretation

MiROvaR is a potential predictor of EOC progression with an independent prognostic impact.

Funding: AIRC and CARIPLO Foundation.

INTRODUCTION

Epithelial ovarian cancer (EOC) is a life-threatening disease characterised by late-stage presentation and high pathological and molecular heterogeneity.[1] Standard treatment for EOC patients is aggressive primary surgery followed by platinum-based chemotherapy. Even for patients who achieve a pathologically complete response, maintaining a disease-free status remains a challenge. In fact, in most of the patients an incurable state of platinum-resistant progressive disease eventually restricts therapeutic options. Despite the impressive t advance in surgical approaches and drug development, EOC patients have experienced little improvement in overall survival in the last 30 years,[2] and the five-year survival rate for advanced-stage patients is still around 30%.[3] Great efforts have been made to develop gene expression-based molecular signatures, but actually few molecular prognostic classifiers have been developed [4-9] even less have been externally validated and no one is clinically available for EOC. One reason lies in the fact that EOC is a genetically plastic disease evolving during progression with an impressive heterogeneity at the time of initial diagnosis. To classify EOC patients for risk of progression, we therefore decided to focus our attention to microRNAs (miRNAs) that act as a master layer of regulation for gene expression and whose number is at least one order of magnitude lower than that of genes. Indeed, despite seminal papers on EOC miRNA profiles have been published [10-12], no data on classification of EOC patients for risk of relapse/progression are at present available. With the assumption that relying on miRNA expression could be a feasible approach to develop a prognostic predictor, we analysed the miRNA expression profiles of 894 EOC samples (to our knowledge, the largest collection so far available) to develop a miRNAs-based predictor of risk of EOC relapse/progression. We further assessed predictor performance and validated its risk-predictive power in two independent case materials.

METHODS

Study design and participants

Three chemo-naive case materials were used for this study.

The training set OC179 was derived from the MITO-2 clinical trial (NCT00326456).[13] Paraffin blocks from the primary tumours of patients enrolled for the MITO-2 trial [13] were provided by 17 out of 43 centres participating to the trial. Of these, 305/549 blocks were excluded for lack of sufficient amounts of tumour for RNA extraction, 18 for the poor quality of extracted RNA, 10 because the tumours derived from metastatic lesions, 30 because they were not chemo-naive (samples were collected at interval debulking surgery after three cycles of chemotherapy), and one for hybridisation failure. After clinical-pathological revision of the available paraffin blocks, RNA extraction, quality control and profiling on human miRNA arrays, 179 cases were eligible for data analysis (Appendix page 2 and 10). The MITO2 sub-population examined in the miRNA study, in comparison with the overall MITO2-trial population, contained a slightly lower number of patients not operated at baseline or with stage IV disease at diagnosis (see Appendix page 4). Accordingly, OC179 showed a longer progression-free survival (PFS) time (22·8 months, 95% CI 18–29 months *vs* 17·1 months, 95% CI 16–19·4) and OS (not yet reached; 95% CI 63–NA *vs* 56·6 months, 95% CI 50–68·2).

Two independent series: the OC263 case material (collected at the Istituto Nazionale dei Tumori - Milan (INT-MI), and at the Centro di Riferimento Oncologico (CRO) Aviano) and the OC452 (from the EOC-TCGA data set https://tcga-data.nci.nih.gov/tcga/ and [5] for clinical data), were identified as validation sets. The OC263 study population was obtained by combining all of the EOC cases profiled by using the Illumina microchip platform at INT-MI, namely the OC130 case material (see [14] for details) and the OC133 case material collected at CRO Aviano. All experimental and clinical information for OC452 and OC130 has been described previously [5,14] and are publicly accessible through the GEO series for OC130 (superseries GSE25204 including GSE25202, GSE25203, and GSE67819) and the TCGA website for OC452 (https://tcga-data.nci.nih.gov/tcga/). For OC133, freshly frozen tumour samples were collected from patients with primary EOC who underwent surgical resection, before any chemotherapeutic treatment, at CRO Aviano. All clinical data and complete follow-up information were available. Tumour staging and grading were in accordance with the International Federation of Gynecology and Obstetrics (FIGO) and the World Health Organisation (WHO) criteria, respectively. A pathologist (VC-Aviano), with specialised expertise in gynaecological pathology reviewed all OC133 pathological data confirming the pathological diagnosis and the required representative percentage of tumor in each sample. None of the tumor samples included in our analysis was macrodissected. Tumor representation among the

three sample collections is similar. The TCGA collection included samples with > 70% of tumor and < 20% necrosis (https://tcga-data.nci.nih.gov/tcga/ and [15]). The same sample characteristics have been adopted for training set OC179 and for the validation set OC263: > 70% of tumor cellularity and < 20% necrosis.

Signed consent was obtained from all of the patients included in the study. For both OC179 and OC263, the investigation was approved by the institutional review boards of participating institutions. In the case of OC263, the study was approved by the Independent Ethics Committee of the INT-MI, where the miRNA profiling was performed. In the case of OC179 derived from the MITO-2 clinical trial,[13] sample collection for translational research purposes was carried out following the approval of a study amendment in 2011. The characteristics of patients included in the study are summarised in Table 1.

Procedures.

The procedures for the RNA extraction and quality controls of OC179 and OC133 (included in the OC263 validation set) and for the miRNA expression profiling of all the case materials analysed, are described in detail in the Appendix page 2 and summarised in Figure 1. The Minimum Information About a Microarray Experiment (MIAME)-compliant data reported in this publication have been deposited in gene expression omnibus of the NCBI,[16] and are accessible through the GEO series using the superseries accession number GSE73583, including GSE73581 (OC179) and GSE73582 (OC133).

The prognostic model was developed using the OC179 data as training set and the OC263 and OC452 data as validation sets. Analyses were performed using R statistical language version 3.1.0 (URL http://www.R-project.org) and R/Bioconductor packages (http://www.bioconductor.org/).

OC263 is a microarray meta-analysis of miRNA microarray data sets generated at INT-MI. The OC263 contains three published data sets, (GSE25202, GSE25203, and GSE67819),[14] listing 130 cases as well as the OC133 data set (GSE73582), that were integrated through the virtualArray R/BioConductor package.[17] The data sets shared the same microarray chip version (Illumina Human_v2 MicroRNA) identifying 1146 miRNAs annotated on miRBase v12.0. The ComBat algorithm [18] was applied to the normalised and log2-transformed data matrices to reduce the likelihood of non-biological technical experimental biases causing batch effects. The resulting integrated data set was named OC263.

OC452 contains the ovarian cancer miRNA microarray profile from the TCGA consortium. The level 1 raw data were downloaded along with the clinical annotations in November 2014 from the TCGA website (https://tcga-data.nci.nih.gov/tcga/). The miRNA expression profiling was performed on Agilent 8 x 15K Human microarrays and identified 799 miRNAs annotated on

miRBase 10. Data were normalised using the robust multiarray average algorithm (RMA). log2-transformed, and filtered using the AgiMicroRna R package as described for the training set (Appendix page 2).

The data pre-processing performed to enable a comparison of the three case materials is described in the Results section.

Progression-free survival (PFS) was defined as the primary clinical endpoint. A semi-supervised prediction method involving principal component analysis developed by Bair and Tibshirani [19] (available through the R package superpc [http://www-stat.stanford.edu/~tibs/superpc]) and already successfully applied to transcriptome data,[20] was adapted to the miRNA expression profiles to identify a prognostic model. Briefly, the significance of each miRNA was measured based on a univariate Cox proportional hazards regression analysis of PFS versus the miRNA log expression level. The miRNA entering into the model were not fixed a priori but were selected on the basis of their FDR (false discovery rate) that defines the expected proportion of false positive results for the balancing of competing demands of sensitivity and specificity, avoiding data overfitting. Imposing a FDR < 0.1 corresponding to $\alpha<0.025$, 35 miRNAs entered into the model then we expect a maximum of 4 miRNAs to be false positives. Subsequently, the principal component analysis was used to reduce the dimensionality of the miRNA included in the model. The first two principal components, capturing 74·2% of miRNA expression variability, were used to obtain a regression coefficient (weight) for each miRNA and develop a model (named MiROvaR) to calculate the prognostic risk score. The OC179 samples were classified as being at low or high risk of progression/relapse after a ten-time cross-validation approach as "internal validation". Risk classification was based on the median index values obtained in the set comprising 90% of the cases (training), with the remaining 10% of the omitted cases (test) classified according to this value. All cases were stratified after reiteration of the entire procedure, omitting a different 10% of cases until each case was excluded. The miRNAs entered into the different cross-validation sets were reported as the percentage of cross-validation support, assessing the percentage in the 10-cross-validated set in which the specific miRNA was selected.

The prognostic risk index for each patient can be computed by the following formula:

$\sum_i w_i x_i + 3{\cdot}196617$,

where w_i and x_i are the weight and logged miRNA expression for the i-th miRNA, respectively. A new sample was predicted as being at high (low) risk if its prognostic index was larger than (smaller than or equal to) 0·07359 which is the median value obtained in cross-validation.

The capability of the model to predict PFS was evaluated through Kaplan–Meier curves and the log-rank test. A 1000-permutations test following a procedure known as "random shuffling" was computed to assess the degree of overfitting in the development of our prognostic model.[21] The

survival data were randomly reassigned among the cases and the entire survival risk prediction process was repeated, assessing the null-distribution of the log-rank test. The tail area of the null-distribution beyond the log-rank statistic of the real data estimates the permutation significance to test the null-hypothesis of the absence of a relationship between PFS and miRNA expression.[22] The permutation test based on 1000 permutations reached P value=0·001.

Performance of MiROvaR was evaluated by ROC curves, details on method are reported in Appendix page 2.

Outcomes

PFS was defined as primary end-point since the main goal of the predictor was to identify early relapsing patients. PFS was defined as the time interval (in months) between the date of random assignment (OC179) or the date of surgery (OC263 and OC452) and the date of progression or death, whichever occurred first, or the date of last follow-up for patients alive without progression. The proportion of cases that relied on date of death for PFS was; 4 out of 124 events for the OC179 training set: 7 out of 195 events for the OC263 validation set 1 and 2 out of 327events for the OC452 validation set 2. Since the training set OC179 derived from the MITO2 clinical trial (see Appendix page 10), for this analysis we used the same PFS definition which started from the date of randomization as an ascertained rule in randomized trials (see Appendix page 3 for details). For OC179 case material, time between randomization and surgery is 1 month (mean, SD=0·46, range 0·23-2·56 months).

Overall survival (OS) was defined as secondary endpoint and was defined as the time (in months) between the date of random assignment (for OC179) or the date of surgery (for validation sets) and the date of death or the date of last contact for surviving patients. Median follow-up times were 73, 44, and 56 months for OC179, OC263, and OC452 case materials, respectively; see Table 1 and details in Appendix page 3.

Statistical analysis

PFS and OS curves were reported according to the Kaplan–Meier method and were compared with the log-rank test. Median estimates, with 95% confidence intervals (CI), were also reported. A Cox univariate model was used to estimate the hazard ratio (HR) for each relevant prognostic variable.

Multivariable analysis using a Cox regression model was used to evaluate the prognostic impact of MiROvaR in the context of concomitant effects of other known prognostic factors (stage and residual disease). The validity of proportional hazards assumption for a Cox model fit has been tested by evaluation of scaled Schoenfeld residuals. The choice of the covariates to be used in the

model was based on several reasons. First, we tried to select, among the variables with known prognostic value, those considered as the stronger ones in terms of PFS prediction in order to have a number of covariates not too high, adequate to the limited sample size of the OC179 training set. Second, we tried to avoid variables with subgroups very small based on current definitions (eg grading). Third, we avoided subjective variables (eg. performance status). Fourth, we avoided variables not available in the validation sets (eg. performance status not available in OC263, histology not informative in the TCGA set where there are only high grade serous tumors). As for age, that in any case is not considered a prognostic factor, its eventual use would not significantly change the result.

Eventually, the covariates that we used for multivariable analyses were FIGO stage and residual disease after primary surgery. Based on the extent of residual disease after primary surgery, the patient population was divided into three groups: no evident residual disease (NED), minimal residual disease (mRD, residual tumour smaller than 1 cm), and gross residual disease (GRD, residual tumour larger than 1 cm). They were then classified into two categories for further analysis: optimal debulking (OD, includes patients NED or with mRD) and suboptimal debulking (SOD, residual tumour larger than 1 cm). We choose the codification OD vs. SOD to be consistent with the paper reporting the MITO2 clinical trial final analysis [13] and to avoid small subgroups that might derive from using a 3-category codification. For both univariate and multivariable analyses, stage and surgical debulking were coded as dichotomous indicator variables (stage III /IV *vs* stage I/II, SOD *vs* OD). However a multivariable analysis for PFS was performed also coding residual disease as a three-levels (NED, mRD, GRD) categorical variable. Patients were then grouped based on similar clinical and pathological characteristics (see Table 1). The chi-square test was used to analyse the distribution of MiROvaR high-risk and MiROvaR low-risk patients in relation to clinical and pathologic variables. A P value <0·05 was considered significant.

All analyses were carried out using R statistical language version 3.1.0 (URL http://www.R-project.org). Graphs were generated using R or GraphPad PRISM (version 5.02) software.

Role of the funding source

No sponsor was involved in the study design, data collection, analysis and interpretation, in writing the manuscript, and in the decision to submit for publication. The following author had access to the raw data: MB, SC, MDM, FP, SP, LDC and DM. The corresponding author had full access to all the data and the final responsibility to submit for publication.

RESULTS

The case materials used to develop the miRNA-based molecular predictor of PFS are summarised in Figure 1. Patients' characteristics and Kaplan–Meier curves are detailed in Table 1 and Appendix page 11, respectively. Overall, 894 EOC cases were analysed at the time of diagnosis. We performed accurate data pre-processing to allow the best comparison among the different platforms and chip arrays. The OC179, OC263, and OC452 miRNA array data were separately filtered to exclude miRNAs which were not detectable in all samples. Data matrices of 921, 706, and 661 miRNAs, respectively were obtained. Since each data set was designed on different miRBase releases, each platform was re-annotated on miRBase release 21·0 (June 2014; http://www.mirbase.org/) at the sequence level. Putative miRNAs, sequences identifying virus miRNAs, non-mature miRNAs, or probe-sets unable to distinguish the members of a miRNA family and those discontinued through different miRBase versions were excluded. A list of 385 unique miRNAs (Appendix page 5) was finally prepared and shared among the platforms and checked using the miRBase Traker tool (www.mirbasetracker.org) [23] to avoid confounding miRNA nomenclature.

On the basis of the defined algorithm and after a 10-time cross validation, we developed a model containing 35 unique miRNAs whose expression significantly contributed to defining the risk of disease progression (Figure 2). Among the 35 identified miRNAs, 16 were associated with better prognosis (putative oncosuppressive miRNAs) and 19 with worse prognosis (putative oncogenic miRNAs) (Figure 2 and Table 2). The 35 miRNAs-based predictor of Risk of Ovarian Cancer Relapse/progression was named MiROvaR.

MiROvaR applied to OC179 clearly separated 89 high- and 90 low-risk patients (HR 1·85, 95% CI 1·29–2·6; P<0·0001) (Figure 3A and Table 3). ROC analyses were used to evaluate MiROvaR performance. The average AUC over 10-time cross-validation reaches a value of 0·68 with an acceptable standard deviation (+/- 0·02) confirming the good performance of our model in OC179 (Appendix page 12). When challenged against validation sets, MiROvaR was able to stratify patients for their risk of progression with significantly different PFS times. One hundred and forty one MiROvaR high-risk (median PFS=12 months, 95% CI 10–13), and 122 MiROvaR low-risk (median PFS=34 months, 95% CI 26–45) patients were identified in the OC263 (Figure 3B) cohort while 283 MiROvaR high-risk (median PFS=15 months, 95% CI 14–18) and 169 MiROvaR low-risk (median PFS=19 months, 95% CI 17–27) patients were detected in the OC452 cohort (Figure 3C). The predictive value of MiROvaR was validated in both case materials: HR 3·16, 95% CI 2·3–4·3, P<0·0001 for OC263 and HR 1·39, 95% CI 1·1–1·74, P=0·0047 for OC452 (Table 3) with AUC = 0·72 (SD ±0·01) and 0·58 (SD ±0·02) in OC263 and OC452, respectively (Appendix page

12). In all three series, advanced stage at diagnosis and sub-optimal debulking after primary surgery were, as expected, significantly associated with progression in univariate analysis (Table 3). Importantly, MiROvaR maintained its independent prognostic impact in all series when analysed in multivariable analysis adjusting for these clinical covariates although its impact on OC452 was less impressive (Table 3 and Appendix page 6). MiROvaR ability to stratify patients' OS in all case materials is reported in Appendix pages 7 and 13. No interactions were observed between MiROvaR and the type of treatment (carboplatin plus taxane *vs* carboplatin plus pegylated doxorubicine) in the OC179 set (p for interaction = 0·62). A significant association of MiROvaR high risk was observed with advanced stage in the OC179 data set and with residual disease in the OC263 series (Appendix page 8).

When compared with the OC452 (EOC-TCGA) validation set selected for high-grade serous ovarian cancer (HGSOC), the OC263 validation set was more heterogeneous in histotype and grading (see Table 1). We confirmed the independent predictive value of MiROvaR by analysing the 230 Type-II (Figure 4A and Table 4), and the 185 HGSOC (Figure 4B and Table 5) cases present in OC263. MiROvaR performance was good also in these two sub-sets of patients with AUC 0·72 (SD ± 0·.02) and 0·71 (SD ± 0·01) for Type-II and HGSOC, respectively (Appendix page 12). The analysis of Type II sub-set was done taking into consideration the new proposed classification of EOC that, besides HGSOC, includes in this sub-set also endometroid high grade, undifferentiated and malignant mixed mullerian tumors [24].

DISCUSSION

Identification of EOC patients with very unfavorable prognosis remains an urgent clinical need to improve the design of tailored therapy. The molecular predictor, MiROvaR, described in this study, and developed on a training set of 179 EOC cases, was able to stratify patients for their risk of relapse/progression with significantly different PFS. The identification of molecular classifiers like MiROvaR is based on an a priori choice of the outcome of interest. Since the main goal of our predictor was to identify early relapsing patients, we defined PFS as the more appropriated end point. PFS is widely accepted as a reasonable end-point in ovarian cancer [25], both clinically and in terms of new drug development, particularly in the first-line of treatment due to the fact that post-progression survival may be quite long and affected also by different and heterogeneous second-line treatments diluting the differences eventually seen in PFS. MiROvaR was able to separate subgroups with different outcome in two independent validation sets, OC263 collected in our Institutions including both frozen and FFPE samples and TCGA relaying on frozen tissues only. We were however able to demonstrate that data obtained on frozen samples could be highly reproduced in FFPE samples and vice versa. [14] Although with a less impressive power, its value was confirmed also in the TCGA data set, the only so far available public collection with fully annotated clinical data that we use as second validation set, thus underlying MiROvaR ability to add significant prognostic information.

Importantly, MiROvaR retained its independent prognostic impact in multivariable analysis. In order to have a number of covariates adequate to the sample size of the OC179 training set, we selected among the known prognostic clinical variables for EOC those considered as the stronger ones in terms of PFS prediction (i.e. FIGO stage and residual disease). MiROvaR showed a good performance despite the criteria adopted for residual disease categorization and even if applied to heterogeneous populations of patients, thus supporting its value in stratifying patients according to their risk of progression, regardless of the clinical–pathological characteristics of their tumours at presentation.. With MiROvaR we aimed to develop a widely useful tool that could encompass the biological/molecular differences among the histological sub-types of ovarian cancers. For this reason we decided not to rely on possibly miRNA-driven patients segregation according to histological characteristic and MiROvaR validation in data sets with different mix of patient characteristics, strengthen its potential use. However, its validity in homogeneous sub-group of patients was confirmed in TCGA dataset that relies on HGSOC only, and in OC263 when considering HGSOC and type II patients which exclude low grade serous ovarian cancer, low grade endometrioid, clear cell and mucinous ovarian cancer. These tumor types (Type I) are rare and poorly represented and would rather benefit of a dedicated study.

The subgroup of patients with a very unfavorable prognosis identified by MiROvaR might be candidate to more aggressive strategies (possible addition of bevacizumab and/or maintenance treatment). However, the ability of more aggressive strategies to improve the prognosis of MiROvaR-high-risk patients should be independently demonstrated before saying that MiROvaR can guide treatment selection. For instance we could not predict response to therapeutic treatments in the analyzed case materials since interactions of potential predictive qualifiers and treatment can only be studied in randomized trials. In this study, OC179, derived from MITO2 randomized clinical trial, represents the only data set where such an analysis could be performed and yielded negative results. Within this dataset there is also a power issue related to the sample size, so we can only say that this analysis does not generate any hypothesis worth of further testing in terms of interaction with treatment arms (conceptually antracycline vs taxane).

From a molecular point of view, one limit of the functional interpretation of miRNA profiles is due to their regulatory role, as each miRNA could regulates numerous genes and the fine tuning of each gene could be different and tissue specific. As specified below, for most of the miRNAs included in MiROvaR, their key role as central nodes in biological processes has been already identified. MiROvaR contains 35 unique miRNAs with an individually different relevance and different impact on patients' prognosis. Among the 16 miRNAs that gave 100% of cross validation support to the MiROvaR predictor, 13 were associated with favourable prognosis and 3 with poor prognosis. Considering the increasing MiROvaR risk of progression, all of the 13 miRNAs individually associated with favourable prognosis and the 3 miRNAs associated with poor prognosis were down-modulated and up-modulated, respectively (see Figure 2). This suggests that the maintenance/loss of potentially oncosuppressive miRNAs has a greater impact on EOC prognosis than the expression/loss of potentially oncogenic miRNAs, in line with the observation that most miRNAs exert an oncosuppressive role and are consequently mostly down-regulated in cancer.[26] Available literature data about their biological role in EOC essentially confirmed our assumptions, although confirmation of the roles of the 3 putative oncogenic miRNAs is limited. Indeed, miR-193a-5p and miR-30b-3p were up-modulated, respectively, in EOC refractory to neo-adjuvant chemotherapy [27] and in low-grade serous EOC compared to fallopian tube.[28] Although miR-29c-5p has been implicated in the regulatory network related to the mesenchymal subtype of HGSOC,[29] information is currently not available concerning its prognostic role in EOC. An oncosuppressive role for miR-29c has, however, been described in colorectal cancer.[30] In contrast, mature data are available for the majority of the 13 putative oncosuppressive miRNAs. In particular, we have previously shown that loss of a ChrXq27.3 miRNA cluster, totally included in the MiROvaR main contributors (miR-508-3p, miR-509-5p, 514a-3p, miR-506-3p, miR-507, miR-509-

3p, miR-513b-5p and miR-513a-5p), is associated with EOC early relapse [14]. This cluster also appears to be down-modulated in the majority of MiROvaR high-risk patients (see Figure 2). A deep functional characterisation of miR-506, a key node of the master miRNA regulatory network related to mesenchymal EOC subtype,[29,31] linked its expression at tumor level [29] to: (i) inhibition of EOC proliferation and induction of senescence,[32] (ii) suppression of the epithelial–mesenchymal transition (EMT),[33] and (iii) an increase in the response to chemotherapy,[34] thus confirming its oncosuppressive role. Besides the ChXq27.3 miRNA cluster, MiROvaR included almost all the members of the miR-200 family (miR-200a, miR-200b, miR-200c, miR-141, and miR-429), which are known EMT regulators,[35,36] as main contributors and loss of miR-200c expression has been associated with relapse even in stage I EOC.[11] Also this miRNA family appears to be down-modulated in the majority of MiROvaR high-risk patients (see Figure 2). Only in the case of miR-592 no data are available concerning its prognostic role in EOC, its expression was, however, predictive of improved outcome in three different cohorts of colorectal cancers [37] thus supporting an oncosuppressive role.

Although definitive data about the biological and prognostic role of all of the miRNAs included in MiROvaR are not yet available in EOC, the main impact on the prediction of early recurrence appears to be associated with EMT regulation. The high number of miRNAs regulating EMT included in our predictor (miR-506 family and miR-200 family) underlines the relevance of cellular reprogramming to a more mesenchymal phenotype as an initiating event during EOC spread and progression. Furthermore, a very recent paper underlined the relevance of loss of miR-200 family members in contributing to recurrent lung metastases after chemotherapy in a breast cancer model, thus suggesting the potential for an EMT-targeting strategy associated with conventional therapy.[38]

By applying a rigorous methodology, we have developed a strong predictor for EOC risk of progression/relapse by using three independent data sets for which mature follow-up data were available. We assessed its performance and validated its potency in two independent data sets with impressive results in the first validation set (OC263). We are aware that our analysis, by relying on the 385 miRNA shared by all the used platforms, may have lost other important miRNAs. However, our work is one of the few attempts in integrating the existing data building a single model that could be fully validated trying to overcome one of the limitation related to miRNA analysis which rely on the use of different platforms and different annotated lists. The applicability of MiROvaR as a useful clinical-grade assay needs further steps following the established framework [39] and guidelines [40] which include: i) the identification of an appropriate methodology (microarray, RTqPCR, Nanostring …); ii) design of specific probes based on the sequences tested in the

microarray chips; iii) validation on independent cohorts of patients with full clinical annotation available.

We remark that our model is completely free to the scientific community and if other pivotal miRNAs will be indentified in future, they can be tested and integrated into MiROvaR. An opportunity to assess MiROvaR performance will be the availability of samples for translational purposes retrospectively collected in the MITO7 clinical trial [41] and prospectively collected in the MITO16 programs (NCT01706120 and NCT01802749); in the latter, tumor collection has become mandatory possibly avoiding attrition bias.

CONTRIBUTORS

All the authors have been involved in this research manuscript. SC, GB, SP, and DM had the idea for the study. MB, SC, LDC, FP and DM designed the experiments. DC, SL, GS, GC, DR centralized and prepared samples from OC179; FR, DL, MLC, MB, SC, and DM collected samples and clinical data for OC130 dataset included into OC263 validation set. GT, EC, RS, VC collected samples and clinical data for OC133 dataset included into OC263 validation set. FR, DL, RS, GS, AS, PS, EB, VM and SP obtained samples and clinical data. SL, GS, MLC, VC, GFZ, MC did the hystopathological analysis. LDC did the microarray analysis and submission of data to GEO. MB, LDC and MDM did the bioinformatic and statistical analyses. MB, SC, DM, LDC, SP and FP critically revised all the data. All the authors reviewed the manuscript and approved the final version.

CONFLICT OF INTEREST

All the authors declare no competing interests.

ACKNOWLEDGMENTS

We thank the Biorepository personnel at the Fondazione IRCCS Istituto Nazionale dei Tumori whose activity made this study possible in particular Dr. Silvia Veneroni for valuable contribution in tissue sample collection. The Italian Association for Cancer Research (AIRC: IG-10302 to SC, IG-17475 to DM, Special Program 12162 to SC and IG-5776 to SP) and CARIPLO Foundation (2013-0865 to DM) for partially supporting this study.

References

1. Jayson GC, Kohn EC, Kitchener HC, Ledermann JA. Ovarian cancer. *Lancet* 2014; **384**: 1376-88.
2. Vaughan S, Coward JI, Bast RC, Jr. *et al*. Rethinking ovarian cancer: recommendations for improving outcomes. *Nat Rev Cancer* 2011; **11**: 719-25.
3. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2015. *CA Cancer J Clin* 2015; **65**: 5-29.
4. Tothill RW, Tinker AV, George J *et al*. Novel molecular subtypes of serous and endometrioid ovarian cancer linked to clinical outcome. *Clin Cancer Res* 2008; **14**: 5198-208.
5. Kang J, D'Andrea AD, Kozono D. A DNA repair pathway-focused score for prediction of outcomes in ovarian cancer treated with platinum-based chemotherapy. *J Natl Cancer Inst* 2012; **104**: 670-81.
6. Tan TZ, Miow QH, Huang RY *et al*. Functional genomics identifies five distinct molecular subtypes with clinical relevance and pathways for growth control in epithelial ovarian cancer. *EMBO Mol Med* 2013; **5**: 983-98.
7. Verhaak RG, Tamayo P, Yang JY *et al*. Prognostically relevant gene signatures of high-grade serous ovarian carcinoma. *J Clin Invest* 2013; **123**: 517-25.
8. Riester M, Wei W, Waldron L *et al*. Risk prediction for late-stage ovarian cancer by meta-analysis of 1525 patient samples. *J Natl Cancer Inst* 2014; **106** (5). pii: dju048.
9. Konecny GE, Wang C, Hamidi H *et al*. Prognostic and therapeutic relevance of molecular subtypes in high-grade serous ovarian cancer. *J Natl Cancer Inst* 2014; **106**: dju249.
10. Iorio MV, Visone R, Di Leva G *et al*. MicroRNA signatures in human ovarian cancer. *Cancer Res* 2007; **67**: 8699-707.
11. Marchini S, Cavalieri D, Fruscio R *et al*. Association between miR-200c and the survival of patients with stage I epithelial ovarian cancer: a retrospective study of two independent tumour tissue collections. *Lancet Oncol* 2011; **12**: 273-85.
12. Vecchione A, Belletti B, Lovat F *et al*. A microRNA signature defines chemoresistance in ovarian cancer through modulation of angiogenesis. *Proc Natl Acad Sci U S A* 2013; **110**: 9845-50.
13. Pignata S, Scambia G, Ferrandina G *et al*. Carboplatin Plus Paclitaxel Versus Carboplatin Plus Pegylated Liposomal Doxorubicin As First-Line Treatment for Patients With Ovarian Cancer: The MITO-2 Randomized Phase III Trial. *J Clin Oncol* 2011; **29**: 3628-35.
14. Bagnoli M, De Cecco L, Granata A *et al*. Identification of a chrXq27.3 microRNA cluster associated with early relapse in advanced stage ovarian cancer patients. *Oncotarget* 2011; **2**: 1265-78.
15. Cancer Genome Atlas Research Network. Integrated genomic analyses of ovarian carcinoma. *Nature* 2011; **474**: 609-15.
16. Edgar R, Domrachev M, Lash AE. Gene Expression Omnibus: NCBI gene expression and hybridization array data repository. *Nucleic Acids Res* 2002; **30**: 207-10.
17. Heider A, Alt R. virtualArray: a R/bioconductor package to merge raw data from different microarray platforms. *BMC Bioinformatics* 2013; **14**: 75.
18. Johnson WE, Li C, Rabinovic A. Adjusting batch effects in microarray expression data using empirical Bayes methods. *Biostatistics* 2007; **8**: 118-27.
19. Bair E, Tibshirani R. Semi-supervised methods to predict patient survival from gene expression data. *PLoS Biol* 2004; **2**: E108.
20. De Cecco L, Bossi P, Locati L, Canevari S, Licitra L. Comprehensive gene expression meta-analysis of head and neck squamous cell carcinoma microarray data defines a robust survival predictor. *Ann Oncol* 2014; **25**: 1628-35.
21. Radmacher MD, McShane LM, Simon R. A paradigm for class prediction using gene expression profiles. *J Comput Biol* 2002; **9**: 505-11.

22. Crijns AP, Fehrmann RS, de JS *et al*. Survival-related profile, pathways, and transcription factors in ovarian cancer. *PLoS Med* 2009; **6**: e24.
23. Van Peer G., Lefever S, Anckaert J *et al*. miRBase Tracker: keeping track of microRNA annotation changes. *Database (Oxford)* 2014; **2014**: bau080.
24. Shih I-M, Kurman RJ. Ovarian tumorigenesis: a proposed model based on morphological and molecular genetic analysis. *Am J Pathol* 2004; **164**: 1511-8.
25. Stuart GC, Kitchener H, Bacon M *et al*. 2010 Gynecologic Cancer InterGroup (GCIG) consensus statement on clinical trials in ovarian cancer: report from the Fourth Ovarian Cancer Consensus Conference. *Int J Gynecol Cancer* 2011; **21**: 750-5.
26. Lu J, Getz G, Miska EA *et al*. MicroRNA expression profiles classify human cancers. *Nature* 2005; **435**: 834-8.
27. Mariani M, McHugh M, Petrillo M *et al*. HGF/c-Met axis drives cancer aggressiveness in the neo-adjuvant setting of ovarian cancer. *Oncotarget* 2014; **5**: 4855-67.
28. Zhang S, Lu Z, Unruh AK *et al*. Clinically relevant microRNAs in ovarian cancer. *Mol Cancer Res* 2015; **13**: 393-401.
29. Yang D, Sun Y, Hu L *et al*. Integrated Analyses Identify a Master MicroRNA Regulatory Network for the Mesenchymal Subtype in Serous Ovarian Cancer. *Cancer Cell* 2013; **23**: 186-99.
30. Zhang JX, Mai SJ, Huang XX *et al*. MiR-29c mediates epithelial-to-mesenchymal transition in human colorectal carcinoma metastasis via PTP4A and GNA13 regulation of beta-catenin signaling. *Ann Oncol* 2014; **25**: 2196-204.
31. Sun Y, Guo F, Bagnoli M *et al*. Key nodes of a microRNA network associated with the integrated mesenchymal subtype of high-grade serous ovarian cancer. *Chin J Cancer* 2015; **34**: 28-40.
32. Liu G, Sun Y, Ji P *et al*. MiR-506 suppresses proliferation and induces senescence by directly targeting the CDK4/6-FOXM1 axis in ovarian cancer. *J Pathol* 2014; **233**: 308-18.
33. Sun Y, Hu L, Zheng H *et al*. MiR-506 inhibits multiple targets in the epithelial-to-mesenchymal transition network and is associated with good prognosis in epithelial ovarian cancer. *J Pathol* 2014; **235**: 25-36.
34. Liu G, Yang D, Rupaimoole R *et al*. Augmentation of response to chemotherapy by microRNA-506 through regulation of RAD51 in serous ovarian cancers. *J Natl Cancer Inst* 2015; **107**: djv108.
35. Park SM, Gaur AB, Lengyel E, Peter ME. The miR-200 family determines the epithelial phenotype of cancer cells by targeting the E-cadherin repressors ZEB1 and ZEB2. *Genes Dev* 2008; **22**: 894-907.
36. Mezzanzanica D, Bagnoli M, De Cecco L, Valeri B, Canevari S. Role of microRNAs in ovarian cancer pathogenesis and potential clinical implications. *Int J Biochem Cell Biol* 2010; **42**: 1262-72.
37. Boisen MK, Dehlendorff C, Linnemann D *et al*. Tissue microRNAs as predictors of outcome in patients with metastatic colorectal cancer treated with first line Capecitabine and Oxaliplatin with or without Bevacizumab. *PLoS One* 2014; **9**: e109430.
38. Fischer KR, Durrans A, Lee S *et al*. Epithelial-to-mesenchymal transition is not required for lung metastasis but contributes to chemoresistance. *Nature* 2015; **527**: 472-6.
39. Micheel CM, Nass SJ, Omenn GS, Editors. *Commitee on the review of Omics-based tests for predicting patient outcomes in clinical trials; Board on health care service; Board on health sciences policy; Institute of medicine; Evolution of translational OMICS: lessons learned and the path foreward.*: The National Academies Press (Washington, D.C.), 2012.
40. Altman DG, McShane LM, Sauerbrei W, Taube SE. Reporting Recommendations for Tumor Marker Prognostic Studies (REMARK): explanation and elaboration. *PLoS Med* 2012; **9**: e1001216.

41. Pignata S, Scambia G, Katsaros D *et al*. Carboplatin plus paclitaxel once a week versus every 3 weeks in patients with advanced ovarian cancer (MITO-7): a randomised, multicentre, open-label, phase 3 trial. *Lancet Oncol* 2014; **15**: 396-405.
42. Callari M, Dugo M, Musella V *et al*. Comparison of microarray platforms for measuring differential microRNA expression in paired normal/cancer colon tissues. *PLoS One* 2012; **7**: e45105.

Table 1.
Clinical and pathological characteristics of patients included in the three case materials

	OC179 (MITO2)		**OC263 (INT-CRO)**		**OC452 (EOC-TCGA)**	
	N° (179)	%	N° (263)	*%*	N° (452)	*%*
Age, years						
mean, median	58, 59		55, 56		59, 58	
range	28-78		25-85		26-87	
Histology						
Serous	124	***69***	190	***72***	452	***100***
Undifferentiated	10	***6***	23	***9***	na	
Endometroid	24	***13***	26	***10***	na	
Mucinous	0	***0***	1	***0***	na	
Clear Cells	6	***3***	7	***3***	na	
Others + Mixed	13	***7***	15	***6***	na	
Missing information	2	***1***	1	***0***	na	
Stage (FIGO)						
I	17	***9***	16	***6***	11	***2***
II	15	***8***	9	***3***	27	***6***
III	123	***69***	212	***81***	350	***77***
IV	24	***13***	26	***10***	63	***14***
Missing Information					1	***0***
Grade						
border line	0	***0***	3	***1***	1	
1, well differentiated	5	***3***	7	***3***	5	***1***
2, moderately differentiated	27	***15***	51	***19***	55	***12***
3, poorly differentiated	126	***70***	177	***67***	382	***84***
Undifferentiated	10	***6***	23	***9***	0	**0**
GX	0	***0***	0	***0***	8	***2***
Missing information	11	***6***	2	***1***	1	
Amount of residual disease						
NED	73	***41***	76	***29***	102	***23***
<1 cm, mRD	42	***23***	85	***32***	208	***46***
>1 cm, GRD	53	***30***	101	***39***	100	***22***
not operated	11	***6***	0	***0***	0	**0**
Missing information	0	***0***	1	***0***	42	***9***
Median follow up (months)	**73 (IQR 60-88)**		**44 (IQR 24-71)**		**56 (IQR 25-86)**	

na=not applicable; FIGO=International federation of Gynecology and Obstetrics; NED=not evident disease; mRD=minimal residual disease; GRD=gross residual disease; IQR=interquartile range.

Table 2
List of the 35 miRNAs entered into the prognostic model

Unique ID	P value	% Cross Validation support	Hazard ratio	0·95 CI	Weight (w_i)
hsa-miR-193a-5p	<0·0001	100	1·977	1·287-2·974	0·010396
hsa-miR-508-3p	<0·0001	100	0·747	0·756-0·958	–0·045965
hsa-miR-509-5p	<0·0001	100	0·684	0·731-0·918	–0·035031
hsa-miR-514a-3p	<0·0001	100	0·811	1·183-2·367	–0·058425
hsa-miR-506-3p	<0·0001	100	0·635	1·031-1·512	–0·032425
hsa-miR-507	<0·0001	100	0·588	1·492-2·612	–0·026022
hsa-miR-509-3p	<0·0001	100	0·783	1·059-2·14	–0·049717
hsa-miR-592	0·00015	100	0·255	1·126-2·356	–0·002782
hsa-miR-29c-5p	0·00071	100	1·595	0·706-0·925	0·005566
hsa-miR-513b-5p	0·00072	100	0·817	0·678-0·911	–0·028496
hsa-miR-513a-5p	0·00074	100	0·766	0·694-0·905	–0·021663
hsa-miR-200c-3p	0·0015	100	0·793	1·181-2·278	–0·027508
hsa-miR-141-3p	0·0017	100	0·819	1·153-2·700	–0·032066
hsa-miR-200b-3p	0·0027	100	0·786	1·232-2·065	–0·028151
hsa-miR-423-5p	0·0029	90	1·765	1·241-3·165	0·005948
hsa-miR-486-5p	0·0030	90	1·345	1·032-1·52	0·015239
hsa-miR-200a-3p	0·0032	100	0·808	1·206-2·854	–0·032221
hsa-miR-23a-5p	0·0052	80	1·641	1·226-2·537	0·006169
hsa-miR-330-3p	0·0061	80	1·856	0·727-0·958	0·004021
hsa-miR-30b-3p	0·0064	100	1·983	1·062-1·531	0·002938
hsa-miR-484	0·0079	80	1·6	1·160-2·206	0·002136
hsa-miR-769-5p	0·0082	70	1·762	1·121-1·612	0·002445
hsa-miR-135b-5p	0·0089	80	0·851	0·479-0·841	–0·024577
hsa-miR-100-3p	0·0090	90	1·958	0·429-0·805	0·003563
hsa-miR-99b-5p	0·0094	70	1·35	0·637-0·874	0·007011
hsa-miR-143-5p	0·0096	80	1·674	0·685-0·895	0·00264
hsa-miR-429	0·012	60	0·835	0·555-0·843	–0·030913
hsa-miR-151a-3p	0·013	60	1·363	0·662-0·886	0·004522
hsa-miR-574-5p	0·016	50	1·283	0·732-0·912	0·005807
hsa-miR-452-5p	0·017	60	1·276	0·726-0·907	0·00919
hsa-miR-29a-5p	0·018	50	1·765	1·049-1·568	0·000855
hsa-miR-195-3p	0·019	40	1·629	0·099-0·661	0·005412
hsa-miR-890	0·023	40	0·085	1·186-2·614	–0·000287
hsa-miR-30d-5p	0·023	40	1·253	0·010-0·717	0·000766
hsa-miR-193b-5p	0·024	60	1·506	1·075-1·695	0·005293

The unique miRNA ID according to miRBase 21·0, p values, percentages of cross validation support, individual hazard ratios, 0·95 Confidence Interval (CI) and weights are reported. miRNAs whose expression is associated with poor prognosis are shown in red, while those whose expression is associated with favourable prognosis are shown in blue.

Table 3
Univariate and multivariable analysis (Cox regression) of progression-free survival for clinical and biological variables in the test set (OC179) and validation sets (OC263 and OC452).

Datasets	Variables	Univariate analysis			Multivariable analysis		
		HR	95% CI	P value	HR	95% CI	P value
OC179 (n=179, events=124)	***Stage***						
	III–IV *vs* I–II	4·74	2·40–9·36	*<0·0001*	3·70	*1·83–7·49*	*<0·00028*
	Surgical debulking						
	SOD *vs* OD	2·10	1·46–3·00	*<0·0001*	1·46	*1·01–2·12*	*0·043*
	miRNA predictor						
	High *vs* Low risk	1·85	1·29–2·64	***<0·00082***	1·48	*1·03–2·13*	***0·036***
OC263 (n=262, events=194)	***Stage***						
	III–IV *vs* I–II	2·16	1·25–3·73	*<0·0057*	*2·16*	*1·21–3·90*	*0·0097*
	Surgical debulking						
	SOD *vs* OD	2·23	1·67–2·97	*<0·0001*	*1·53*	*1·13–2·08*	*0·0060*
	miRNA predictor						
	High *vs* Low risk	3·16	2·33–4·29	***<0·0001***	*3·09*	*2·24–4·28*	***<0·0001***
OC452 (n=409, events=300)	***Stage***						
	III–IV *vs* I–II	1·68	1·02–2·78	*0·04*	*1·79*	*1·04–3·08*	*0·035*
	Surgical debulking						
	SOD *vs* OD	1·37	1·07–1·75	*0·012*	*1·27*	*0·99-1·63*	*0·059*
	miRNA predictor						
	High *vs* Low risk	1·39	1·11–1·74	***0·0047***	*1·41*	*1·11-1·79*	***0·0047***

HR=hazard ratio; CI=confidence interval; OD=optimal debulking; SOD=suboptimal debulking.

Table 4
Univariate and multivariable analysis (Cox regression) of progression-free survival for clinical and biological variables in Type-II cases (n=230) of OC263 case materials

	Variables	Univariate analysis			Multivariable analysis		
		HR	**95% CI**	**P value**	**HR**	**95% CI**	**P value**
***OC263* (n=230, events=172)**	***Stage***						
	III–IV *vs* I–II	2·45	1·20–5·00	*0·013*	*2·37*	*1·10–5·12*	*0·028*
	Surgical debulking						
	SOD *vs* OD	2·07	1·53–2·81	*<0·0001*	*1·50*	*1·10–2·06*	*0·011*
	miRNA predictor						
	High *vs* Low risk	3·25	2·34–4·51	*<0·0001*	*3·16*	*2·24–4·45*	*<0·0001*

HR=hazard ratio; CI=confidence interval; OD=optimal debulking; SOD=suboptimal debulking.

Table 5
Univariate and multivariable analysis (Cox regression) of progression-free survival for clinical and biological variables in HGSOC cases (n=185) of OC263 case materials

	Variables	Univariate analysis			Multivariable analysis		
		HR	**95% CI**	**P value**	**HR**	**95% CI**	**P value**
***OC263* (n=185, events= 140)**	***Stage***						
	III–IV *vs* I–II	2·81	1·15–6·90	*0·023*	*2·67*	*0·96–7·38*	*0·058*
	Surgical debulking						
	SOD *vs* OD	2·10	1·50–2·95	*<0·0001*	*1·62*	*1·14–2·29*	*0·0071*
	miRNA predictor						
	High *vs* Low risk	3·00	2·09–4·3	*<0·0001*	*2·96*	*2·03–4·31*	*<0·0001*

HR=hazard ratio; CI=confidence interval; OD=optimal debulking; SOD=suboptimal debulking.

FIGURE LEGENDS

Figure 1. Characteristics of the case materials, miRNA platforms and Chip arrays used for development of a miRNA classifier able to predict EOC patients' risk of relapse. FFPE = formalin-fixed paraffin embedded samples. The inter-platform reproducibility of miRNA microarray profiles was demonstrated by our previous study by Callari et al. [42].

Figure 2. Expression heat map of the 35 miRNAs entering into the predictive model. Columns = patients (179) and rows = miRNAs (35), sorted on the basis of the established index. The plot above summarises the specific MiROvaR risk-score index for each sample. Blue: miRNAs whose expression is associated with a good prognosis. Red: miRNAs whose expression is associated with a poor prognosis.

Figure 3. Kaplan–Meier curves of patients stratified according to the miRNA predictor. **A.** MiROvaR stratification of patients included in the OC179 training set. MiROvaR high-risk (red line) and MiROvaR low-risk (blue line) curves were compared using a log-rank test. Nyr = not yet reached. **B, C.** The ability of MiROvaR to classify EOC patients for risk of progression was validated in two independent data sets. Kaplan–Meier curves of OC263 (B) and OC452 (C) patients stratified according to the miRNA predictor are shown. Blue lines = MiROvaR low-risk patients; red lines = MiROvaR high-risk patients.

Figure 4. Ability of MiROvaR to predict the risk of progression in Type-II (**A**) and high-grade serous ovarian cancer (HGSOC), **B**) subpopulations of the OC263 validation set. Blue lines = MiROvaR low-risk patients; red lines = MiROvaR high-risk patients.

ID		Reference	type of material	n° samples	miRNA platform	impact on study
OC179 (MITO2)		Present study	FFPE	179	Agilent miR-BASE 17	training set
OC263 (INT-CRO)	OC130	(14)	frozen/FFPE	130	Illumina miR-BASE 12	validation set 1
	OC133	Present study	frozen	133		
OC452 (EOC-TGCA)		(5)	frozen	452	Agilent miR-BASE 10	validation set 2

Figure 1

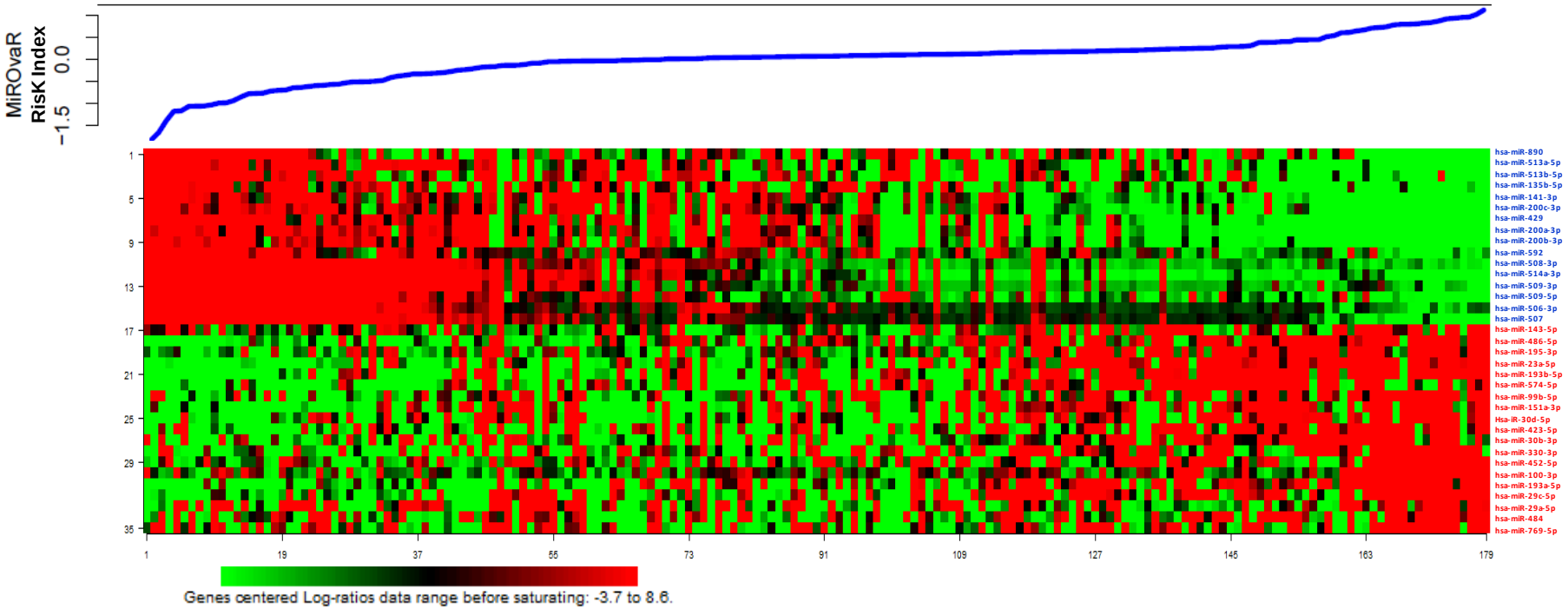
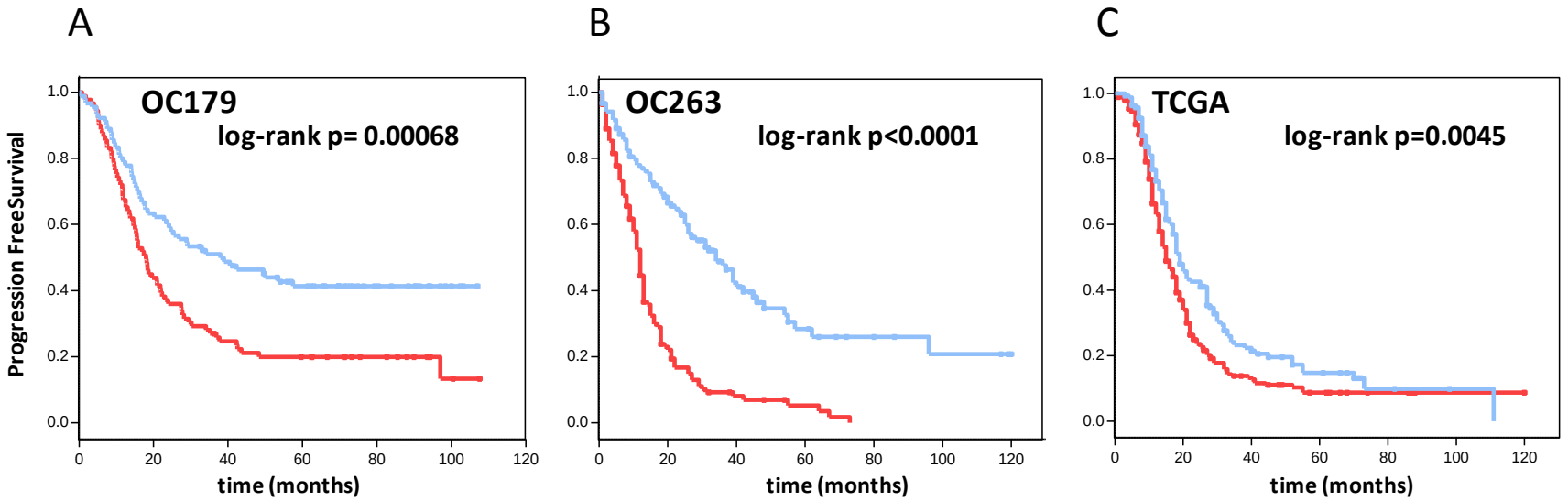

MiROvaR
RisK Index
−1.5
0.0
1
5
9
13
17
21
25
29
35
1
19
37
55
73
91
109
127
145
163
179
hsa-miR-890
hsa-miR-513a-5p
hsa-miR-513b-5p
hsa-miR-135b-5p
hsa-miR-141-3p
hsa-miR-200c-3p
hsa-miR-429
hsa-miR-200a-3p
hsa-miR-200b-3p
hsa-miR-592
hsa-miR-508-3p
hsa-miR-514a-3p
hsa-miR-509-3p
hsa-miR-509-5p
hsa-miR-506-3p
hsa-miR-507
hsa-miR-143-5p
hsa-miR-486-5p
hsa-miR-195-3p
hsa-miR-23a-5p
hsa-miR-193b-5p
hsa-miR-574-5p
hsa-miR-99b-5p
hsa-miR-151a-3p
Hsa-iR-30d-5p
hsa-miR-423-5p
hsa-miR-30b-3p
hsa-miR-330-3p
hsa-miR-452-5p
hsa-miR-100-3p
hsa-miR-193a-5p
hsa-miR-29c-5p
hsa-miR-29a-5p
hsa-miR-484
hsa-miR-769-5p
Genes centered Log-ratios data range before saturating: -3.7 to 8.6.

Figure 2



Number at risk

Low	90	58	43	31	15	4	1	121	77	32	14	8	5	2	169	62	25	12	4	2	0
High	89	40	22	17	10	3	1	141	26	8	4	0	0	0	283	83	24	11	6	2	2

Censored

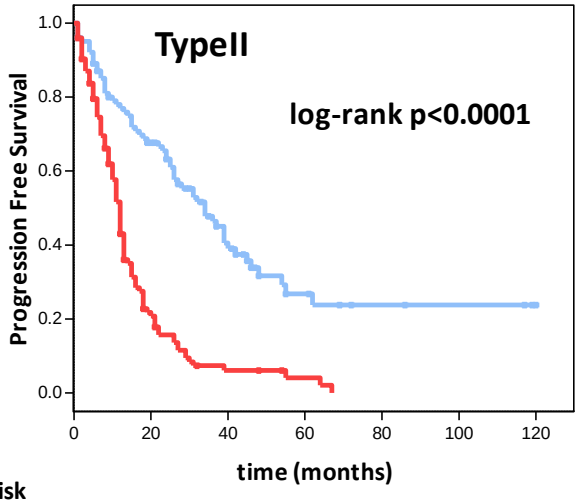
Low	0	0	2	8	24	35	38	0	9	27	38	43	45	49	0	33	40	45	52	54	54
High	0	0	1	2	9	15	17	0	14	17	19	19	19	19	0	46	54	61	67	70	71

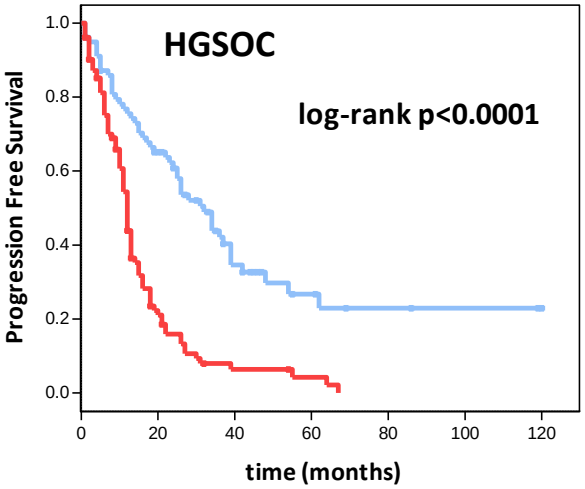
Risk prediction	N° patients	Events	Median PFS	95%CI
Low	90	52	38	24-nyr
High	89	72	18	15-22

Risk prediction	N° patients	Events	Median PFS	95%CI
Low	122	73	34	26-45
High	141	122	12	10-13

Risk prediction	N° patients	Events	Median PFS	95%CI
Low	169	115	19	17-27
High	283	212	15	14-18

Figure 3 A-C





Number at risk

Low	102	65	27	12	7	6	3	79	49	21	9	6	6	2
High	128	22	6	3	0	0	0	106	18	5	3	0	0	0

Censored

Low	0	7	21	31	34	36	40	0	5	15	22	24	26	29
High	0	13	15	17	17	17	17	0	12	14	15	15	15	15

Risk prediction	N° patients	Events	Median PFS	95%CI
Low	102	62	34	26-42
High	128	111	12	10-13

Risk prediction	N° patients	Events	Median PFS	95%CI
Low	79	50	32	25-39
High	106	91	12	11-13

Figure 4 A-B